DOCUMENT RESUME

ED 057 095 TM 000 939

AUTHOR Masonis, Edward J.
TITLE Problems and Procedures in Planning a Situation Based Video Test on Teaching.
INSTITUTION Educational Testing Service, Princeton, N.J.
PUB DATE Apr 71
NOTE 8p.; Paper presented at the Annual Meeting of the American Personnel and Guidance Association, Atlantic City, New Jersey, April 1971

EDRS PRICE MF-$0.65 HC-$3.29
DESCRIPTORS Audiovisual Programs; Planning; *Situational Tests; Teacher Behavior; Teacher Education; *Teacher Evaluation; Teaching Skills; *Television; *Test Construction; *Testing Problems
IDENTIFIERS *Project NuTEx

ABSTRACT

This paper briefly outlines some problems one must solve when developing a video-based test to evaluate what a teacher knows about learning and instruction. Consideration is given to the effect the use of videotapes of actual classroom behavior have on test planning. Two methods of incorporating such situational material into the test specifications are discussed, and some perceived advantages and disadvantages of each are presented. An account is given of the planning method employed in Project NuTEx, a research project supported by the National Teacher Examinations, along with some evaluative statements about the effects the adopted method had on the project. The many ways of presenting objective test items in a television test are described, and the problems this increased flexibility presents are briefly outlined. Some indication is made of research questions that must be answered if television is to be used rationally as a testing medium. (Author/AG)

ED057095

JUL 3 1971

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY

PROBLEMS AND PROCEDURES IN PLANNING A SITUATION BASED VIDEO TEST ON TEACHING

by

Edward J. Masonis
Program Director
Administrative and Supervisory
Examination Services

Educational Testing Service

April 1971

TM 000 939

For a while now, the educational community has been anticipating the time when man's knowledge of technology is combined with his knowledge of testing and evaluation. This combination seems to be a natural, especially in light of the generally pronounced limitations of the standard paper and pencil test, and the apparent possibilities technology offers to eliminate these limitations. Consider, for example, a paper and pencil test on teaching. When the test maker decides to construct a test of this kind, he usually includes items that attempt to measure the student's ability to apply the principles of teaching and learning, and when he does, he is almost certainly allowing his test to be open to criticism. This criticism arises because these so-called application items seldom provide the student with an adequate base on which to choose the correct response. Those of you who have submitted to a test on teaching, guidance, or administration may have experienced the frustration of attempting to answer an item of this kind. You read the stem and the options and realize that, depending on the details of a situation, any of the options could be correct; therefore, you must begin to search for the answer the test maker had in mind.

A possible solution to this problem was planned by several people at Educational Testing Service. Their notion was to investigate the possibility of using television as a means of improving tests on teaching.

Since I was one of several staff members at ETS involved in the actual production of a prototype television test of this kind, it would be

comforting for me to be able to say that our efforts were a complete success and that we are now ready to apply our newly-acquired knowledge toward building new versions of the National Teacher Examinations or other tests in which the use of television might be appropriate. Unfortunately, we were not totally successful. Like all efforts of this kind, we experienced some bad luck along the way that helped to keep us from reaching our goal, and because we had few guidelines to help make decisions about how to proceed, some of our decisions were incorrect. We did learn a great deal, however, and we are convinced that in the near future testing of this kind will be sufficiently refined for use in a national testing program such as the NTE. With this thought in mind, I would like to share with you some of the procedures we used as well as some of the problems we encountered as we proceeded to work on the project.

Since my basic involvement in the project was limited to the more conventional aspects of planning and producing the test, and did not include the production of the taped situations used as a basis for the test, I will limit my remarks to a description of the test planning phase of the project. Originally, I intended to discuss the planning and specification problems a test developer encounters in putting together both literal video based tests, i.e., a video test using words only and no pictures, and a situational video test, i.e., a video test using taped segments of actual or staged behavior. Because of time pressures, however, I will limit my presentation to a discussion of some basic planning concepts

and problems one encounters in attempting to build a situation-based, behavior-oriented video test. The development of a video test of this kind obviously requires the test developer to solve more problems than must be solved when a standard paper and pencil test is being developed. Consider, first of all, the initial planning stage in the test development process. The development of a standard paper and pencil test begins with the construction of a set of specifications that minimally include a statement of the content to be tested, as well as a statement of the mental operations or skills the student is expected to perform in answering the test items. These specifications are usually developed by outlining the content first, and then identifying the skills. It doesn't necessarily have to be that way, it just seems easier for most test developers to follow that pattern.

In contrast, a situation-based video test requires additional planning. The most obvious additional information is a need to describe the kind of situational material that is to be used as the basis for the test questions. At this stage, the test developer is confronted with a chicken or egg problem as he must decide which comes first, the content or the situations. He may decide to identify the kinds of situations to use as the basis for the test, and then attempt to identify the concepts and skills that can be tested from these situations, or, he may decide to identify the content and skills to be tested, and then describe situations to fit the content and operations specifications. However, to accept either of these methods also requires the acceptance of certain advantages and limitations that are inherent in each.

If the test developer uses the first method and begins by identifying the situations on which the test will be based, he has the advantage of not having to describe the situational behavior very specifically. He may, in fact, employ very general terms in his description, or perhaps he may decide not to describe the situations at all. Thus, he may use unrehearsed, real-life, situations rather than the staged variety, and be able to capitalize on the spontaneity of real life behavior. Unfortunately, this method requires the test developer to accept serious limitations on the content to be tested, since he can only ask questions about the content that is related to or a part of the behavior presented in the situation. Since the validity and reliability of a test are affected by the number and kinds of items a test contains, and since the concepts being tested are considered by most educators to be the essential part of the test, this procedure may be considered by many to be impractical.

If the test developer decides to use the second technique in planning the test, that is, begin by identifying the skills and concepts and then describing the situations needed, he has the advantage of having more flexibility in specifying the concepts to be tested, but he is faced with the problem of limiting the kinds of situations he can include. In fact, he may find it impossible to use real-life situations because the situations needed to test the concepts outlined in the specifications would be impossible to locate in advance. Therefore, he may have to use staged situations. The resulting artificiality of the staged scenes may seriously detract from the test to a point where some may consider this plan equally impractical.

In planning the situation-based video items for Project Nutex, we decided to have the television crew go to public schools and tape actual classroom behavior. We relied on the crew's ability to produce tapes in sufficient quantity and quality to use as the basis for the test. In effect, the content and process axes of the test specifications were determined by the behavior recorded on the videotapes.

Again, it would be most pleasing for me to report that our decision was correct, for once it was made, it shaped our efforts throughout the project. Unfortunately, our decision proved to be a handicap in test production; and near the end of our year's work, I concluded that it would have been better to proceed by developing the content and process dimensions of the specifications first, and then describing the kinds of situations needed to evaluate the content, and produce the situations by some sort of role-playing technique. I felt that the loss in "reality" resulting from this staging technique would be more than compensated for by a desirable gain in the overall quality of the test itself. Put another way, the method we adopted made it easier to obtain situation tapes but encouraged the violation of some essential principles of measurement. The other method would make the production of situation tapes more difficult and perhaps more artificial, but it would allow a person to employ sound test development procedures.

An examination of some of our situation based items illustrates the shortcomings of the test development procedure we adopted. Because the test

development crew had no control over the content of the taped situations, they had to frame items about the concepts and principles related to the behaviors recorded on the tape. As it happened, the television crew focused primarily on teacher behavior to the neglect of student behavior, and the teachers tended to rely heavily on telling or information-giving behavior while they were being taped. These circumstances forced the test development crew to frame items on only a few dimensions of classroom behavior, e.q., information-giving technique, questioning technique, classroom atmosphere, and concept teaching patterns.

Aside from the problems a test developer incurs while attempting to coordinate the content and processes axes of the test specifications wtih the content of the videotaped situations, another planning consideration in preparing a video based test needs to be mentioned. Our experience with the production of an all literal video test illustrated that a video-based test provides the test developer with much more flexibility in presenting the test items than is afforded by a paper and pencil test. To illustrate, the test items in a multiple choice video test can be presented on the screen just as it might appear in a printed booklet, i.e., with the stem and all the options presented together, or the stem may be presented first, and each option may be presented under the stem in an additive fashion, or the stem may be presented and each option may be separately flashed under the stem as it is read by the announcer. Many other patterns can be used, and these deal only with the test item and not its relationship to the video taped

situation on which it will be based. When this relationship is also considered, many more possibilities for presenting the items occur. The increased flexibility provided by these alternatives requires more careful consideration of the presentation format in the planning stage of test development. The test developer is forced to consider what item format will best measure the concept he wants measured on the cognitive level he prescribes with the situation being used as the basis for the item. If he does not, he may plan and produce a test that does not fully capitalize on the special advantages the video mode contributes to the testing situation.